

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

(19)日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平8-98146

(43)公開日 平成8年(1996)4月12日

(51)  $\text{InLCl}^8$

識別記号

戸内整理番号

FI

技術表示箇所

H04N 5/93

G 1 0 L 9/18

G 1 1 B 20/10

A

E 7736-5D

H04N 5/93

G

審査請求 未請求 請求項の数6 O.L (全 10 頁)

(21)出願番号 特願平6-229129

(22)出題日 平成6年(1994)9月26日

(71)出願人 000004226

日本電信電話株式会社

東京都新宿区西新宿三丁目19番2号

(72) 發明者 森 達男

東京都千代田区内幸町一丁目1番6号 日

本電信電話株式会社内

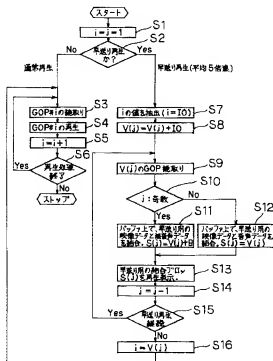
(74) 代理人 弁理士 若林 忠

(54) 【発明の名称】 スキップサーチ方法

(57) 【要約】

【目的】 高速感がある映像と意味が分かる音声とを有する再生が可能な早送りのためのスキップサーチ方法の提供。

【構成】 映像と音声を含む連続画像のフレームh個からなるGOPを一定数m個を連続して群とし、複数の群を映像・音声フィールドとして群単位に同期して格納しておき、倍速度に応じて群毎に映像用のGOPを一定の位置にスキップして選択して読み出し、さらに、音声用のGOPの意味が分かる再生が可能のように映像用のGOPの一部を無音声化し、または、ノイズ化し、選択された映像用のGOPの映像ブロックと音声ブロックが残されたGOPの音声ブロック、あるいは、音声のみのブロックを結合して再生する。



## 【特許請求の範囲】

【請求項1】 連続映像から同期をとって所定の単位長毎に分割された映像および音声のファイルを平均  $n$  倍速で再生する際、前記単位長の  $m$  倍を一群とした群毎に  $k = m/n$  個の単位長の映像を選択し、群毎に  $k$  個以下の単位長の音声を選択し結合して再生するスキップサーチ方法において、

音声ファイルについて群内で選択される単位長の数が  $k$  より少ない数値である  $p$  個であり、 $k - p$  個の無音またはノイズの単位長と音声ファイルの  $p$  個の単位長を結合して

【請求項2】 各群内における映像ファイルを選択される単位長相互間の間隔が一定間隔にならないように選択する請求項1記載のスキップサーチ方法。

【請求項3】 各群内で映像ファイルについて選択される単位長相互間の間隔が漸増した後に漸減するように選択する請求項1記載のスキップサーチ方法。

【請求項4】 前記無音またはノイズとした映像の単位長と音声ファイルの単位長の区間を交互に結合する請求項1記載のスキップサーチ方法。

【請求項5】 前記群毎に  $m$  個の単位長の中から連続した、 $k$  個より小さい  $p$  個の音声の単位長を選出し、 $k$  個の映像の単位長と該  $p$  個の音声の単位長と  $k - p$  個の無音またはノイズの単位長を結合して再生する請求項1記載のスキップサーチ方法。

【請求項6】 前記所定の単位長がグループオブピクチャ単位である請求項1乃至5のいずれか1項に記載のスキップサーチ方法。

## 【発明の詳細な説明】

## 【0001】

【産業上の利用分野】本発明は、リクエスト型映像・音声情報提供システムセンタ等において、蓄積装置にMP E G 符号化方式等でデジタル圧縮した映像・音声を書き込み、その読み取りを行う映像・音声情報の蓄積読み取り方法に関し、特に音声出力を可能とした早送り再生方法に関する。

## 【0002】

【従来の技術】従来、リクエスト型映像・音声情報提供システムでは、センタと多数の端末とを接続し、端末から要求された映像・音声情報をセンタの蓄積装置から読み取って当該端末に送信していた。また、従来のビデオの符号化はフレームを基本として行われていたが、圧縮率向上のためのフレーム間予測符号化では早送りや逆転再生等の機能を実現できなかった。フレーム内符号化画面を挟み込んだ、複数のフレームを集めてグループオブピクチャ（以下GOPと称す）を構成している。早送りを実現するためには、（1）映像GOPを一定間隔でスキップして再生するスキップサーチ方法、（2）フレーム内符号化圧縮されたフレームのみを選んで再生する早

送り方法等がある。

## 【0003】

【発明が解決しようとする課題】上述した従来のスキップサーチ方法は、（1）の一定間隔でGOPをスキップして再生する方法では通常再生となる映像再生部分が長くなるので高速感がなくなり、逆に連続したGOPが少ないと、実質的な通常再生となる部分が短くなり、音声の意味が不明瞭となる。また、（2）のフレーム内符号化画面のみを選んで早送り再生する方法では、フレームを飛び飛びに再生するため、再生フレーム対応の短い音声飛び飛びに出力され、音声の意味が不明瞭になるという欠陥があった。

【0004】本発明の目的は、高速感がある映像と意味がある音声とを併せて再生することを可能とする音声付映像スキップサーチ方法を提供することである。

## 【0005】

【課題を解決するための手段】本発明のスキップサーチ方法は、連続映像から同期をとって所定の単位長毎に分割された映像および音声のファイルを平均  $n$  倍速で再生する際、前記単位長の  $m$  倍を一群とした群毎に  $k = m/n$  個の単位長の映像を選択し、前記群毎に  $k$  個以下の単位長の音声を選択し結合して再生するスキップサーチ方法において、音声ファイルについて群内で選択される単位長の数が  $k$  より少ない数値である  $p$  個であり、 $k - p$  個の無音またはノイズの単位長と音声ファイルの  $p$  個の単位長を結合して音声と再生する処理を行う。

【0006】また、各群内における映像ファイルを選択される単位長相互間の間隔が一定間隔にならないように選択する処理、および、各群内で映像ファイルについて選択される単位長相互間の間隔が漸増した後に漸減するように選択する処理、さらに、前記無音またはノイズとした映像の単位長と音声ファイルの単位長の区間を交互に結合する処理も本発明に含まれる。

【0007】また、群毎に  $m$  個の単位長の中から連続した、 $k$  個より小さい  $p$  個の音声の単位長を選出し、 $k$  個の映像の単位長と該  $p$  個の音声の単位長と  $k - p$  個の無音の単位長を結合して再生する処理も本発明の望ましい実施態様である。

【0008】前記所定の単位長はグループオブピクチャ単位であってもよい。

## 【0009】

【作用】デジタルの映像・音声ファイルを複数のGOPに分割し、かつ、GOPをさらに  $h$  個（ $h \geq 2$  を満足する整数）の複数のフレームで構成し、映像と音声をフレーム単位に同期させて蓄積しておき、 $m$  個のGOPを1群とし、群毎に  $k$  個の映像再生用GOPを選出し、選出した  $k$  個のGOPの識別子が奇数となるGOPの音声ブロックを無音声データ、または、ノイズで置換し、あるいは、 $m$  個のGOPの中から連続した  $p$  個のGOPを音声再生用を選択してその音声ブロックを映像ブロック

と結合して再生するので、意味が分かる音声出力を含む映像再生が可能となる。

#### 【0010】

【実施例】次に、本発明の実施例について図面を参照して説明する。

【0011】図1は本発明の連続映像のGOPとフレームとの構成関係を示す図、図2(a)は5倍速再生のGOP読み取り選択順位を示す図、図2(b)は図2

(a)によって選択したGOPの映像と音声の結合を示す図、図3は本発明のスキップサーチ方法の第1の実施例のフローチャートである。

【0012】本実施例における連続映像のデジタル動画情報(映像および音声情報)ファイルは、図1に示す15の映像フレーム構成のGOP m個(本実施例ではm=20)を一群として任意の数の群を蓄積装置の物理的に連続する領域に蓄積する。図中GOPはG001のようにGOPの識別子GとGOPに付与した通し番号のGOP番号(3桁)とで表示し、フレームは同様にf001、f002のようにフレームの識別子fと3桁の通し番号で表示する。さらに本実施例では一群のGOPの再生表示のための選択数kをk=4とし、群構成は連続したGOP番号のGOPをm=20個づつ区切ったものとする。また、選出順序が奇数番目のGOPの音声部分を無音声データまたはノイズで置き換えて、音声と無音声データが交互に、あるいは、音声とノイズが交互に並ぶようにする。以下、音声を加加させた平均5倍速(n=\*

$$a_{i+1} = a_i + d_{i+1} = a_i + c$$

#### 【0017】

##### 【数1】

$$\sum_{j=1}^k a_j = m - k \quad (12)$$

$$a_j \geq 1$$

式11、12より

#### 【0018】

$$\sum_{j=1}^k a_j = m - k = k \times a_1 + c \times \sum_{j=1}^{k-1} k = k \times a_1 + c \times (k-1) \times k + 2$$

$$\therefore m = (1 + a_1 + c \times (k-1) \div 2) \times k \quad (14)$$

ここで、m=20、k=4とした平均5倍速では、式★40★(14)は、

$$20 = 4 + 4 \times a_1 + 6 \times c$$

$$\therefore 2 \times a_1 + 3 \times c = 8$$

となる。図2(a)は、a<sub>1</sub>=1、c=2、m=20、k=4の場合の平均5倍速モードの場合のGOPの読み取り図である。また、図2(b)は、平均5倍速モードにおける映像・音声結合図である。図2(a)において、i番目に読み取る映像データ格納GOPの番号V

$$V(i) = (m-k) \left[ (i-1) \div k \right] + V + (i-1 - [(i-1) \div k] \times k) \div 16 \left[ (i-1) \div \right.$$

\* m/k=5)の変速型映像スキップサーチ方法を説明する。

【0013】第1の実施例はスキップ幅が漸増するモードである。

【0014】また、本実施例の動画情報ファイルは、動画情報が6ビット/秒のビットレートで符号化されたビット列とし、30フレーム/秒の動画情報を有するものとする。GOP1個には15フレームを格納し、GOP4個単位に読み取るものとする、1フレームの表示時間は1/30秒であるから、4GOP分すなわち60フレームでは60/30秒=2秒となる。したがって、4GOPの読み取り時間(Tr)は表示時間(2秒)より小さくなければならない。蓄積装置として、アクセス時間50m秒、読み取り速度20Mビット/秒の磁気ディスク装置を例とすれば、GOP4個分(2秒相当)を読み取り処理する所要時間は、Tr=50m秒+(6Mビット/20Mビット/秒)×2=0.65秒となり、2秒以下という条件を十分満たしている。本実施例の6Mビット/秒の動画情報(音声情報も含む)では3多重(2÷0.65≒3)までの読み取り・再生表示が可能となる。

【0015】本実施例の具体例を以下に示す。

【0016】一つの群内のスキップモードとしてスキップ幅a<sub>i</sub>が漸増するモードの実施例である。スキップGOP数の増加分(d<sub>i+1</sub>)を一定とし、cとすると、

$$(11)$$

#### ※【数2】

$$4] + i + (i-1 - [(i-1) \div 4] \times 4) \div 2$$

また、スキップサーチ時に再生出力される音声と映像との結合ブロックをS(i)とすると、S(i)は、iが偶数のときは音声ブロックA(i)を使用し、iが奇数のときには無音声データ(蓄積装置の専用エリアBに予め格納しておく)を使用する。したがって、

$$S(i) = V(i) - B \quad (i \text{ は奇数})$$

$$S(i) = V(i) - A(i) \quad (i \text{ は偶数})$$

となる。

【0019】図3は本発明の第1の実施例のGOP読み取り・再生のフローチャートである。図中Snはステップ番号示す。

【0020】1) まず、ファイル内のGOPポインタ*i*および早送りGOP数カウンタ*j*を1に初期設定( $i=j=1$ )する(ステップ1)。

【0021】2) 再生種別の指示が早送り再生か通常再生かを判定する(ステップ2)。

3) 通常再生のときは、ポインタ*i*のGOPを読み取る(ステップ3)。

【0022】4) 読み取ったGOPポインタのGOPを再生する(ステップ4)。

【0023】5) 次に、GOPポインタをインクリメントする( $i=i+1$ ) (ステップ5)。

【0024】6) 予め指定されているGOPポインタまで再生処理が達したかどうかを確かめて未達の場合はステップ3へ処理を回し、終了を確認すると処理を停止する(ステップ6)。

【0025】7) 一方、早送り再生(高速再生)をステップ2で判定すると早送り再生開始のGOPポインタの値( $i=j=1$ )を抽出する(ステップ7)。

【0026】8) 早送り再生の読み取りGOPのアドレスを算出する(ステップ8)。具体的には  $V(j) = V(j) + 10$  となる。

【0027】9)  $V(j)$  のGOPを読み取る(ステップ9)。

10)  $j$  が奇数か偶数かを判定する(ステップ10)。

11)  $j$  が奇数の場合は、バッファ上で  $V(j)$  のGOPの映像ブロックと  $A(j)$  のGOPの音声ブロック

$$a_j \geq 1$$

上記の式21、22より

【0032】

$$\sum_{j=1}^k a_j = m - k = k \times a_1 + d \times \sum_{j=1}^k k = k \times a_1 + d \times (k-1) \times k \div 2$$

$$\therefore m = (1 + a_1 + d \times (k-1) \div 2) \times k \quad (24)$$

ここで、 $m=20$ 、 $k=4$ とした平均5倍速の場合の例★ ★では、24式は、

$$20 = 4 - 4 \times a_1 + 6 \times d$$

$$\therefore 2a_1 + 3d = 8$$

$$(25)$$

となる。図6(a)は  $a_1=1$ 、 $d=2$ 、 $m=20$ 、 $k=4$ の場合の平均5倍速モードにおけるGOP読み取り図、図6(b)は平均5倍速モードの映像・音声結合図である。図6(a)において、 $i$ 番目に読み取る映像ブロックがあるGOPの番号を  $V(i)$  とし、 $|$  を絶対値記号、 $□$  をガウス記号とすると、 $V(i) = e^i + \{5 - 2e - 3\} / 2$  但、 $e = i - 4 \times [(i-1) \div 4]$  となる。

【0033】早送り再生時の音声と映像との結合プロ

\* クを結合する(ステップ11)。すなわち

$$S(j) = V(j) - A(j) = V(j) - B$$

ここで、 $B$ は無音データの格納ブロックである。

12) ステップ10で  $j$  が偶数の場合は、バッファ上で、映像ブロックと音声ブロックとを結合する(ステップ12)。すなわち

$$S(j) = V(j) + A(j) = V(j)$$

13) 早送り再生用データとして、 $S(j)$  を再生表示する(ステップ13)。

14) 早送り再生の出力ポインタ  $j$  をインクリメントする( $j=j+1$ ) (ステップ14)。

15) 早送り再生が予め設定されている群に達したかどうかにより再生処理を継続するかどうかを判定し、継続の場合はステップ9へ処理を回し(ステップ15)、16) 終了の場合はステップ3に処理を移して通常再生に戻す(ステップ16)。

【0028】図4はG100のGOPまで通常再生を行い、G101から図3に示す早送り再生を行った場合の平均5倍速の映像・音声出力図、図5は平均5倍速の信

速変化図である。

【0029】次に、第2の実施例について説明する。

【0030】第2の実施例は、スキップ幅が漸増した後に漸減する場合である。ここで、スキップGOP数の増加分( $d_{in}$ )は、増加する場合は  $d$  (一定) とし、ピークに達して漸減する場合には  $-d$  (一定) とすると、 $a_{in} = a_1 + d_{in}$  (21)

【0031】

【数3】

$$\sum_{j=1}^k a_j = m - k \quad (22)$$

$$(23)$$

※【数4】

クを  $S(i)$  とし、 $S(i)$  には、 $i$  が偶数のときは音声データ  $A(i)$  を使用し、 $i$  が奇数のときは蓄積装置の専用エリア  $B$  に格納した無音データを使用する。これにより、 $U$  を理論和の記号とすると、

$$S(i) = V(i) \cup A(i) \quad (i \text{ は偶数})$$

$$S(i) = V(i) \cup B \quad (i \text{ は奇数})$$

となる。図7はG100のGOPまでの通常再生を行い、G101よりスキップ幅が漸増した後漸減するモードで早送り再生を行った場合の平均5倍速の映像と音

声の出力図、図8は平均5倍速の倍速変化図である。

【0034】次に、第3の実施例について説明する。

【0035】本実施例は連続するGOPを $m=20$ 個づつの群構成とし、群毎の再生表示GOP数を $k=4$ 、群毎の連続音声のGOP数を $p=3$ とすると、 $i$ 番目に読み取るべき映像ブロックがあるGOPの番号 $V(i)$ は以下のとおりとなる。ここに $\#$ を $i$ 番目示す記号とすると、

$$V(i) = \#(m-k) \left[ (i-1) \div k \right] + i + \left[ (i-1) \div k \right] \times k$$

図9(a)は、 $a_1=1$ 、 $m=20$ 、 $k=4$ 、 $p=3$ の場合における平均5倍速モードのGOP読み取り図で、この数値を代入すると、 $V(i) = \#16 \left[ (i-1) \div 4 \right] + 1 + (i-1) \div 4 \times 4$  となる。また、 $i$ 番目に読み取るべき音声ブロックがあるGOP番号を $A(i)$ とすると、 $i$ が4の倍数のときは、 $A(i) = \#B$  (但し  $i=4n$ ) の位置に格納されている無音データを読み取る。 $i$ が4の倍数でないときは、\*

$$\begin{aligned} S(1) &= V(1) \cup A(1) = \#01 \cup \#01 = \#01 \\ S(2) &= V(2) \cup A(2) = \#03 \cup \#02 \\ S(3) &= V(3) \cup A(3) = \#07 \cup \#03 \\ S(4) &= V(4) \cup A(4) = \#13 \cup \#B \end{aligned}$$

で表現される。図9(b)は、平均5倍速における映像・音声結合図で図中の◎記号は無音となる音声データBを配置していることを示す。

【0038】

【発明の効果】以上説明するように本発明は、群毎に $k$ 個の映像早送り再生用のGOPを抽出し、かつ、抽出GOPの識別子が奇数となるGOPの音声ブロックを無音またはノイズとし、あるいは、抽出したGOPの中の何れかを連続した $p$ 個のGOPの音声ブロックを映像ブロックと結合して再生するので、以下の効果がある。

【0039】1. 連続音声を付加した映像のスキップサーチなので、映像シーンの意味が分かり易くなる。

【0040】2. GOP単位のサーチに対してGOPのスキップ幅、あるいは倍速度をダイナミックに変化させているので、早送りの速度感を出すことができる。

【0041】3. GOP単位の音声と音声のつなぎ部分に無音データ、または、ノイズを挟み込んでいるので、音声が滑らかに聞こえる。

【0042】以上の効果により、MPEG方式等のフレーム間符合理化方式によるGOP構成の映像番組のピクチャサーチに効果的である。

$$\begin{aligned} *A(i) &= \#m \times \left[ (i-1) \div k \right] + i = 20 \times \left[ (i-1) \div 4 \right] + i \\ &\quad (\text{但し } i \neq 4n) \text{ の位置に格納されている。} V(1), \end{aligned}$$

および $A(i)$ は、

$$\begin{aligned} V(1) &= \#01, & A(1) &= \#01 \\ V(2) &= \#03, & A(2) &= \#02 \\ V(3) &= \#07, & A(3) &= \#03 \\ V(4) &= \#13, & A(4) &= \#B \end{aligned}$$

となる。

【0036】ここで、 $V(1)=A(1)=\#01$ 、 $V(2)=A(3)=\#03$  が重複するので、 $V$

(1)、 $A(2)$ 、 $V(3)$ 、 $V(4)$ 、 $A(4)$ の数値に相当するGOP番号のGOPを読み取る。

【0037】スキップサーチ用の映像と音声の結合ブロックを $S(i)$ とし、 $U$ を論理和の記号とすると、

$$\begin{aligned} p \geq i \geq 1 & \quad S(i) = V(i) \cup A(i) \\ k \geq j \geq p+1 & \quad S(i) = V(i) \cup B \end{aligned}$$

となる。第1周期を例にとると、

【図面の簡単な説明】

【図1】本発明に使用されているGOPの構成を示す図である。

【図2】(a)は本発明のスキップサーチ方法の第1の実施例のGOP読み取りを示す図、(b)は読み取ったGOPの映像・音声結合図である。

【図3】本発明のスキップサーチ方法の第1の実施例のフローチャートである。

【図4】第1の実施例の平均5倍速とする映像・音声出力図である。

【図5】第1の実施例における映像の平均5倍速の倍速変化によるGOP読み取り状態を示す図である。

【図6】(a)は第3の実施例の平均5倍速によるGOP読み取り図、(b)はその映像・音声結合図である。

【図7】第3の実施例の平均5倍速による映像・音声出力図である。

【図8】第3の実施例の平均5倍速のためのGOP読み出し状態を示す図である。

【図9】(a)は第3の実施例の平均5倍速のためのGOP読み取り図、(b)はその映像・音声結合図である。

【図1】

5001	1001	1002	1003	1004	1005	1006	1007	1008	1009	1010	1011	1012	1013	1014	1015
------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

5002	1016	1017	1018	1019	1020	1021	1022	1023	1024	1025	1026	1027	1028	1029	1030
------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

⋮

【図2】

第1 群	GOP 番号	01	03				07						13							
映像	V01		V03				V07						V13							
音声	—		A03				—						A13							

第2 群	GOP 番号	21	23				27						33							
映像	V21		V23				V27						V33							
音声	—		A23				—						A33							

第3 群	GOP 番号	41	43				47						53							
映像	V41		V43				V47						V53							
音声	—		A43				—						A53							

【図4】

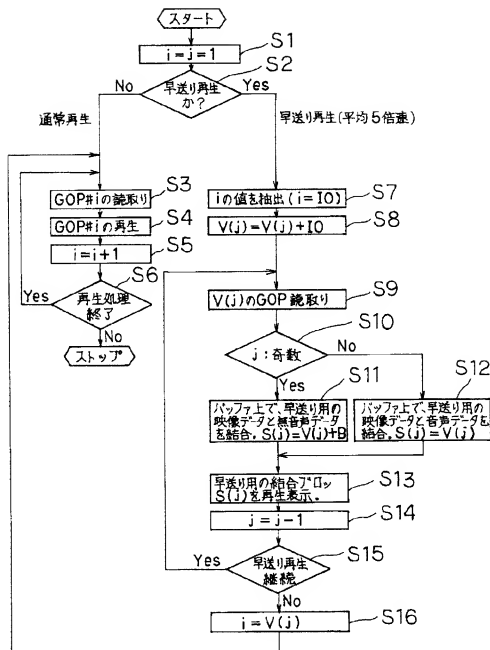
	再生方向→					→→→					→→→				
	通常再生					早送り再生開始					早送り再生①				
											早送り再生②				
映像→	B096	B097	B098	B099	B100	V101	V103	V107	V113	V121	V123	V127	V133		
音声→	A096	A097	A098	A099	A100	@	A103	@	A113	@	A123	@	A133		
						(無音)					(無音)				

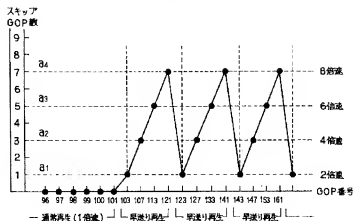
	→→→					→→→					→→→				
	早送り再生③					早送り再生④					早送り再生⑤				
映像→	V141	V143	V147	V153	V161	V163	V167	V173	V181	V183	V187	V193			
音声→	@	A143	@	A153	@	A163	@	A173	@	A183	@	A193			
	(無音)					(無音)					(無音)				



【図3】



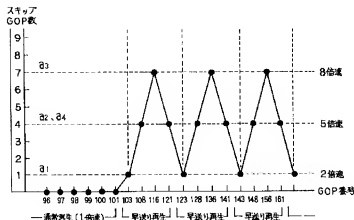
【図5】



【図7】



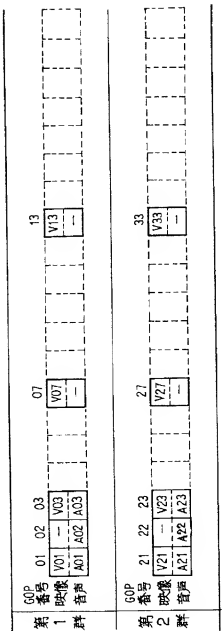
【図8】





【図9】

(a)



(b)

